DOCUMENT RESUME

ED 216 909 SE 037 894

AUTHOR	Kahl, Stuart R.; And Others
TITLE	Sex-Related Differences in Pre-College Science: Findings of the Science Meta-Analysis Project.
INSTITUTION	Colorado Univ., Boulder.
SPONS AGENCY	National Science Foundation, Washington, D.C.
PUB DATE	Mar 82
NOTE	18p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, March 18-23, 1982).
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Academic Achievement; Cognitive Processes; Elementary School Science; Elementary Secondary Education; Females; *Mathematics Education; Physical Sciences; Science Education; *Secondary School Science; *Sex Differences; *Student Characteristics
IDENTIFIERS	Mathematics Education Research; *Meta Analysis; National Assessment of Educational Progress; National Science Foundation; *Science Education Research

ABSTRACT

This project examined science education research relating to relative effects of different science programs, teaching systems, instructional methodologies, teacher preparation programs, and relationships between various student outcomes and student/teacher relationships. More than 80 studies were examined providing information on sex-related differences in pre-college science. Throughout these years, males on the average outperform females in science achievement. On the surface, these sex differences seem small (little more than a tenth of a standard deviation at the senior high level). However, when cognitive outcome differences are broken down by cognitive process levels and by specific science disciplines, some differences are considerably larger. At the higher cognitive levels, the difference favoring senior high school males is as high as a fifth of a standard deviation. In the physical sciences, the male advantage is as high as a third of a standard deviation. Thus, with respect to important stepping stones to post-secondary programs and ultimately careers in science, women are at a disadvantage, not because of statistically significant differences between the sexes, but rather because of the educationally significant magnitudes of those differences. Problems encountered in studies of sex differences in mathematics education are also addressed. (Author/JN)

SEX-RELATED DIFFERENCES IN PRE-COLLEGE SCIENCE: FINDINGS OF THE SCIENCE META-ANALYSIS PROJECT*

Stuart R. Kahl
RMC Research Corp.

M. Lynette Fleming
University of Colorado

Mark R. Malone
University of Colorado

March 1982

* Paper presented at the annual meeting of the American Educational Research Association, New York, March 1982. For additional information regarding the NSF-funded Science Meta-Analysis Project, contact: Dr. Ronald D. Anderson, Project Director, School of Education, University of Colorado, Boulder, Colorado 80309.

SEX-RELATED DIFFERENCES IN PRE-COLLEGE SCIENCE: FINDINGS OF THE SCIENCE META-ANALYSIS PROJECT

ABSTRACT

The Science Meta-Analysis Project, funded in 1980 by NSF, examined a sizable portion of the research in science education. The meta-analysis questions studied were unusually broad, relating to the relative effects of different science programs, teaching systems, specific instructional methodologies and teacher education programs as well as the relationships between a wide range of student outcomes and teacher and student characteristics.

More than eighty studies were examined which provided information specific to the issue of sex-related differences in pre-college science. Throughout the pre-college years, male students on the average outperform female students in science achievement. On the surface, these differences between the sexes seem small -- little more than a tenth of a standard deviation at the senior high level However, when cognitive outcome differences are broken down by cognitive process levels and by specific science disciplines, some differences are considerably larger. At the higher cognitive levels, the difference favoring senior high school males is as high as a fifth of a standard deviation. In the physical sciences, the male advantage is as high as a third of a standard deviation. Thus, with respect to important stepping stones to post-secondary programs and ultimately careers in science, women are at a disadvantage -- not because of statistically significant differences between the sexes, but rather because of the educationally significant magnitudes of those differences.

The Science Meta-Analysis Project did not analyze research intended to <u>explain</u> sex-related differences in science achievement. In that area of research, science educators seem to be following in the footsteps of researchers in mathematics education. The author cautions science educators of the pitfalls the latter group has encountered in studies of sex-related differences in mathematics. (author)

SEX-RELATED DIFFERENCES IN PRE-COLLEGE SCIENCE:
FINDINGS OF THE SCIENCE META-ANALYSIS PROJECT

In the fall of 1980, the National Science Foundation funded the Science Meta-Analysis Project. The purpose of the study was to consolidate the findings of a large body of research in science education using the technique of meta-analysis to seek answers to some very broad and important research questions. One of the seven major questions pertained to the relationships between student characteristics and science-related student outcomes. Although concerned with a wide variety of student characteristics, the researchers tackling this question encountered a large number of studies which examined sex-related differences in science. The major purpose of this paper is to summarize the findings of that one portion of the overall project.

The past decade has seen considerable growth in the attention given to sex-related differences in learning, particularly in mathematics and science. A primary stimulus, of course, has been the underrepresentation of women in traditionally male-dominated areas of study and work. Logically researchers have investigated differences in the preparation of males and females for entry into these fields. Such differences pertain to both cognitive and affective variables as well as course enrollments. Thus, a great deal of information about the magnitude of these sex-related differences has become available. An obvious next step for researchers is to synthesize this information.

The Larger Study

The Science Meta-Analysis Project was somewhat unique in that it called for the application of the technique of meta-analysis on a grand scale in the investigation of some very general questions. They were:

1. What are the effects of different curricular programs in pre-college science (e.g., ESS, SCIS, S-APA, ISCS, IPS, ESCP)?
2. What are the effects of different instructional systems used in science teaching (e.g., self-paced, computer-assisted, individualized, programmed, tutorial)?
3. What are the effects of different teaching techniques (e.g., questioning behaviors, testing practices, inductive/deductive approaches, manipulatives, demonstration)?

4. What are the effects of the nature and structure of content on student outcomes (e.g., advanced organizers, kinetic structure, cognitive level, sequencing)?

5. What are the effects of different pre- and in-service science teacher training programs and practices (e.g., competency-based, field vs. university-based, summer institutes, content emphasis, training in specific techniques)?

6. What are the relationships between teacher characteristics/behaviors and student outcomes in science?

7. What are the relationships between student characteristics and student outcomes in science?

The effects of interest pertained to a wide variety of student cognitive and affective outcomes in pre-college science. Each of the above questions was investigated by a researcher or team of researchers from one of seven universities across the country. The study was coordinated by the project staff at the University of Colorado under the direction of Dr. Ronald Anderson. Because of the scope of the meta-analysis questions, the Colorado staff, with the cooperation of the ERIC center in Columbus, Ohio, assisted all of the research teams in the literature search process focusing on post-1960 research. Initial emphasis was on dissertations readily available on microfilm from the ERIC center where personnel also assisted in later computer and hand searches of other literature sources. Concentrating on dissertations first reduced the volume of published articles to be utilized in the meta-analysis as many of them are based on dissertation research. Contrary to the opinion of some, dissertation studies (at least those of use in meta-analysis) do not appear to be of lesser quality generally than other research. That would not be a problem in meta-analysis anyway. Furthermore, there is the possibility of bias in the results of studies researchers choose to publish. Certainly the reporting in dissertations is much more complete than that in journal articles, and therefore information important for meta-analysis is less likely to be omitted.

The Technique of Meta-Analysis

Early in the project, the researchers from the various sites were brought to the University of Colorado to be trained in the methodology of meta-analysis by Dr. Gene Glass who devised the technique. The participants then returned to their home institutions to conduct their respective meta-analyses, with Colorado staff available by telephone to assist the researchers when they encountered difficulties.

For each relevant study in a meta-analysis, at least one coding form (much like a very detailed questionnaire) is completed. Of primary concern is the "effect size" from each study. For many research questions such as the first five listed in the previous section, it is useful to determine an effect size or "delta" which is the difference between two groups on some outcome measure in standard deviation units. For example, in evaluating the effectiveness of innovative curricular programs, one might define delta by $\Delta=(\bar{X}_I-\bar{X}_T)/s_T$ where $\bar{X}_I$ is the mean score on the dependent variable for the group using the innovative program, $\bar{X}_T$ the corresponding mean for the group receiving traditional instruction, and s_T the standard deviation of the "traditional" group. There could be instances when the standard deviation of either group or an average of the two would be desirable; however, there is frequently some logic to using traditional or control groups' standard deviations since they would more accurately represent the general situation. Thus, the delta is a measure of the magnitude of a treatment effect in standardized units. If a study measures several different outcome variables, then of course several deltas could be computed for that study.

A great deal more information would be recorded on the coding sheets -- namely more detailed information on the characteristics of the treatments, the sample, the outcome variables and their measures, and the research itself. The primary analysis is often nothing more than the computation of average effect sizes for various categories of effects. Thus, the researchers investigating the first question of concern in the project could determine the average effects on various outcomes of the SCIS program, or inquiry-oriented programs, or inquiry-oriented programs at the secondary level. More sophisticated analyses are appropriate in some instances.

There are a great number of statistical techniques utilized in research, and therefore a variety of statistics are reported. Frequently the "ingredients" for computing deltas directly are not provided in a research report, but fortunately the concept of a delta is so basic that the statistics which are reported are easily transformed into deltas. Occasionally some assumptions have to be made when information required by a transformation is missing, but the delta seems to be extremely robust with respect to such assumptions.

For some questions such as the sixth and seventh project questions which pertain to relationships between teacher or student characteristics and student outcomes, zero-order correlations are more appropriate for meta-analysis than deltas. They can be analyzed in the same manner as deltas.

The technique of meta-analysis has been viewed somewhat suspiciously by a few individuals. Glass et al (1981) and Glass (1982) deal effectively with specific criticisms leveled against the technique. Generally, its exploratory nature, the irrelevance of the significance levels from the studies used, and the inappropriateness of significance testing of the meta-analysis results themselves seem to go against a widespread, but narrow view of the nature of quantitative research. One must remember that significance levels are not indicative of magnitudes of effects, the latter being more important if educational significance is a concern. Unfortunately, it seems that it is only for purposes of meta-analysis that effect sizes are even computed! Treating effect sizes from studies as individuals in analyses may seem strange to some people. It makes sense, however, when one considers how and why effect sizes would be distributed, whether or not they are statistioally significant. Measurement theory treats test items like individuals. To do the same with effect sizes does not constitute a very great leap. Surely an average delta based on several studies is a better indicator of the magnitude of an effect than an effect size based on one study with all of its idiosyncracies, including the particular measures used reflecting one person's definitions of variables.

The Study of Student Characteristics

One team of researchers studied the relationships between student characteristics and student outcomes in science. Student characteristics of interest included (1) various affective variables such as attitude toward school and toward science, anxiety, motivation, and self-concept (some of these were also treated as outcome variables in relation to other characteristics); (2) academic variables such as verbal and quantitative abilities, science course backgrounds, spatial skills, and cognitive level of development; and (3) personological variables such as age, sex, race and socioeconomic status. Outcome variables included such things as attitudes toward science, scientists and science class; achievement in science broken down where possible by cognitive process levels such as knowledge, understanding and higher processes; and miscellaneous variables such as critical thinking, decision-making skills, and science process skills.

Although three or four hundred research reports were reviewed, 169 were actually coded and utilized in the meta-analysis. These studies produced

308 effects in the form of correlations. For sex and race, the corresponding deltas were also computed since deltas seem more meaningful for these two independent variables. Coded studies included 122 dissertations, 41 journal articles, 5 unpublished ERIC documents, and the 1976-77 National Assessment in Science. A large number of relevant journal articles were not used because the studies had already been coded as dissertations.

Approximately half of the studies coded provided information on sex-related differences. Some reports provided by personnel from other project sites gave results of analyses in which sex was used as an independent variable in factorial designs. These too were used to determine effect sizes for sex. In all, almost 150 such effect sizes were computed.

Both correlations between sex and student outcomes and deltas representing standardized sex-related differences in outcomes were recorded so that positive values corresponded to differences favoring males. The deltas reported in this paper are based on standard deviations which are derived from the average variances of the males and females. Thus, $\Delta = \bar{X}_M - \bar{X}_F / s_{avg}$.

While more detailed information pertaining to this component of the project is available (Fleming and Malone, 1982), the general approach of meta-analysis is such that merely reporting the results describes the methodology, provided one understands the nature of an effect -- either a delta or a correlation. One matter, however, does warrant discussion before results are presented -- the weighting of effect sizes in the compucation of average effects.

Frequently the question is asked, "Shouldn't average effect sizes be weighted averages based on study effects weighted by the numbers of subjects involved in the studies?" The arguments for a negative response are most convincing. Such weighting would be like computing a class average on a test using scores weighted by the actual weights in pounds of the individuals. It is the whole individual who achieves a certain performance level, not every one-pound unit of flesh and bone. In meta-analysis, each study is unique in many ways, even if there is commonality in the relationships being investigated by the studies. To assign one study a greater weight would attach greater importance to the variable definitions and measurement instruments used in that study -- a matter unrelated to sample size, the basis of such weights! Unweighted effect sizes value each researcher's definitions and instruments equally. Variability in study results is surely due more to these

unique aspects of studies than to sample size.

A weighting issue which should be of greater concern pertains to situations in which several effect sizes are derived from a single study. For example, if one study yields several sex-related differences (deltas) in mathematics comprehension because several measures of mathematics comprehension were used, it would be best to compute an average delta for that variable for that study, so that the study does not dominate the averages across studies. Multiple effects from a study, however, would not be a problem if they are used in different aggregations. Thus, a single study can contribute a delta in the computation of an average delta for mathematics comprehension and another delta in the computation of an average delta for mathematics anxiety.

In the study reported herein, a large number of effect sizes were based on data from the Third National Assessment in Science. Instead of consolidating effects from National Assessment in the manner described above, average effects are reported separately for all studies, NAEP only, and non-NAEP only.

Sex-Related Differences -- Results

In order to compare the influence of sex on science-related outcomes with that of other factors, it is useful to consider effect sizes in terms of zero-order correlation coefficients. (Keep in mind the variable sex should be interpreted broadly, encompassing not only biological sex, but a vast number of societal/environmental influences as well. In other words, the correlations do not represent the relationships between genetic differences specifically and science outcomes.) Table 1 reports the average correlations between various student characteristics and cognitive and affective student outcomes. The most obvious conclusion to be drawn from this table is that sex has a relatively weak relationship to science outcomes when compared to other factors. Of the personological variables, race and socioeconomic status (SES) have considerably more influence than sex with respect to cognitive outcomes in particular. In terms of group differences, this finding simply means that sex-related differences are very small compared to race differences or differences among SES groups. While this finding tells us nothing new, it is important that the magnitude of sex-related differences at pre-college levels be kept in perspective. Table 1 also documents the

TABLE 1

Average Correlations between Student Characteristics and Cognitive and Affective Student Outcomes

Student Characteristic		Elementary Cog.	Elementary Aff.	Junior High Cog.	Junior High Aff.	Senior High Cog.	Senior High Aff.
Age	$\bar{r}$	.30	.15	.42	.15	.01	-.02
	n	16	2	7	1	9	3
Race* (white/black)	$\bar{r}$	.17	.03	.19	.01	.15	-.02
	n	13	3	12	4	10	4
Race* (white/hispanic)	$\bar{r}$	.13	.02	.09	.01	.06	.02
	n	12	3	10	4	10	4
Socioeconomic Status	$\bar{r}$	.30	.09	.29	.02	.28	.00
	n	19	3	13	5	14	5
Sex	$\bar{r}$	.05	.10	.08	-.01	.07	.07
	n	41	11	25	7	45	19
Attitude/Motivation (school/science)	$\bar{r}$	.31		.19		.34	
	n	4		3		6	
Self Concept	$\bar{r}$	.07		.35		.19	
	n	1		4		4	
Internality	$\bar{r}$	.43		.62		.52	.29
	n	2		1		1	2
Math Ability/ Performance	$\bar{r}$	.46	.04	.52	-.16	.45	.39
	n	8	1	3	1	15	1
Spatial Ability	$\bar{r}$					.44	
	n					5	
Verbal/Language Ability	$\bar{r}$	.55	.07	.59		.47	
	n	13	1	3		8	
Reading Ability	$\bar{r}$	.35	.05	.62	-.04	.43	.39
	n	11	1	5	1	5	1
IQ	$\bar{r}$	.42	.19	.43	.12	.46	.21
	n	27	3	14	5	19	3
Study Skills	$\bar{r}$	.54		.46		.48	.52
	n	5		2		2	1
Cognitive Level of Devel.	$\bar{r}$	.53		.61		.50	-.17
	n	5		1		1	1
Science Course Background	$\bar{r}$					.19	.58
	n					6	1

* Race comparisons based almost entirely on NAEP results.

obvious importance of students' attitudes and aptitudes in relation to science outcomes.

Table 2 shows average sex-related differences in standard deviation units for cognitive and affective outcomes. For elementary school students research results are quite consistent for cognitive outcomes as indicated by the relatively small standard deviation of the deltas (s=.17). Differences are practically nonexistent, the average delta giving males the edge by only one-twentieth of a standard deviation approximately. However, with respect to affective variables, males at the elementary level do appear to have a greater advantage, one of nearly one-fifth of a standard deviation.

Research findings pertaining to cognitive outcomes for junior high males and females are considerably more variable (s=.35). On the average, males outperform females by almost a quarter of a standard deviation on cognitive measures. Based of a small number of studies, there is the suggestion that females may enjoy an advantage with respect to affective outcomes at the junior high level. However, the few deltas are quite variable.

The results of studies involving senior high school students are quite consistent. Male students have an advantage of just over a tenth of a standard deviation on both cognitive and affective outcomes. To put that into perspective, if both kinds of outcomes were measured in IQ-like units, males would average 100.9 and females 99.1.

Average deltas reported in Table 2 for affective outcomes in elementary school and for both cognitive and affective outcomes in junior high show discrepancies between NAEP and non-NAEP results. One should view the NAEP science assessment as just another study in a meta-analysis for reasons explained in the previous section. There are discrepancies just as great among the non-NAEP deltas as between the NAEP and non-NAEP deltas. The limited number of deltas for junior high affective outcomes should be more of a concern, especially considering the variability of results.

In Table 3, deltas are reported for finer breakdowns of cognitive outcomes. Keeping in mind the limited number of deltas in "fine cells", it does appear that compared to the general results reported in Table 2, the male advantage is somewhat greater with respect to science comprehension and application in the elementary years; application in junior high; and comprehension, application and higher processes at the senior high level. There is little sex-related difference in science process skills at any level with the possible exception of senior high.

TABLE 2

Average Standardized Sex-Related Differences (Deltas*)

in Cognitive and Affective Student Outcomes

			All	NAEP	non-NAEP
Elementary					
	Cognitive	$\bar{\Delta}$	.06	.05	.06
		s	.17	.08	.19
		n	36	10	26
	Affective	$\bar{\Delta}$	.18	.04	.25
		s	.25	.05	.28
		n	9	3	6
Junior High					
	Cognitive	$\bar{\Delta}$	.23	.07	.36
		s	.35	.07	.44
		n	22	10	12
	Affective	$\bar{\Delta}$	-.11	.10	-.40
		s	.37	.06	.44
		n	7	4	3
Senior High					
	Cognitive	$\bar{\Delta}$	.12	.12	.12
		s	.24	.08	.28
		n	37	10	27
	Affective	$\bar{\Delta}$	.12	.11	.12
		s	.13	.14	.14
		n	15	4	11

* $\Delta = \bar{X}_M - \bar{X}_F / s$ where s is based on the average of the male and female variances on the outcome measures.

TABLE 3

Average Sex-Related Differences (Deltas) in Selected Cognitive Outcomes in Science

Cognitive Outcome		Elementary			Junior High			Senior High		
		All	NAEP	non-NAEP	All	NAEP	non-NAEP	All	NAEP	non-NAEP
Knowledge	$\bar{\Delta}$	.05	.10	.05	.09	.09		.10	.13	.07
	s	.25	0	.27	0	0		.04	0	0
	n	9	1	8	1	1		2	1	1
Comprehension	$\bar{\Delta}$	.16	.09	.22	.16	.10	.19	.20	.14	.22
	s	.09	0	0	.05	0	.01	.07	0	.07
	n	2	1	1	3	1	2	4	1	3
Application	$\bar{\Delta}$	.17	.05	.28	.30	.08	.51	.19	.13	.21
	s	.16	0	0	.30	0	0	.13	0	.16
	n	2	1	1	2	1	1	4	1	3
Higher Processes	$\bar{\Delta}$	-.01	.01	-.03	.16	.05	.26	.15	.08	.22
	s	.03	0	0	.15	0	0	.10	0	0
	n	2	1	1	2	1	1	2	1	1
General (not enough info. to sort as above)	$\bar{\Delta}$	.04	.05	.04	.32	.11	.45	.15	.15	.14
	s	.15	.12	.16	.47	.07	.56	.27	.07	.31
	n	16	4	12	11	4	7	17	4	13
Science Process Skills	$\bar{\Delta}$	.05	.03	.06	.03	.03		.07	.18	.02
	s	.10	0	.13	0	0		.11	0	.08
	n	3	1	2	1	1		3	1	2

Table 4 shows the average deltas for cognitive outcomes associated with selected science disciplines at the secondary school levels. There appears to be little difference between the performance levels of males and females in biology and earth science. However, males have a decided advantage in other disciplines -- general science ($\bar{\Delta}$=.29), physical science ($\bar{\Delta}$=.33), chemistry ($\bar{\Delta}$=.16) and physics ($\bar{\Delta}$=.22).

Discussion

The Magnitude of Differences

Studies of sex-related differences in science are nothing new. For years, we have been hearing of studies which found or failed to find significant differences between the sexes in science achievement and attitudes. Unfortunately, researchers have tended to ignore the magnitudes of those differences and seem to have been more interested in statistical significance. Studies with little statistical power have led to conclusions that there are probably no such differences even when the differences they actually obtained were reasonably large. Conversely, some studies with considerable power have resulted in claims that there probably are sex-related differences in science based on relatively small differences. Most people, upon reviewing the literature piecemeal, would probably conclude the latter, that there probably are differences, but would be hard-pressed to estimate the magnitude of them. Fortunately, the technique of meta-analysis provides a way to determine a best estimate of the differences.

The findings summarized in the previous section give us some cause for concern. Throughout the pre-college years, male students on the average outperform female students in science achievement. Males also have the advantage with respect to science-related attitudes at the elementary and senior high levels. On the surface, these differences between the sexes seem small and perhaps of little concern. At the senior high level, the male advantage on cognitive and affective variables is only a little more than a tenth of a standard deviation. However, when cognitive outcome differences are broken down by cognitive process levels and by specific science disciplines, some differences are considerably larger. At the higher cognitive levels, the difference favoring senior high males is as high as a fifth of a standard deviation. In the physical sciences the male advantage is as high as a third

TABLE 4

Average Sex-Related Differences (Deltas) in Cognitive Science Outcomes for Selected Science Disciplines (Secondary Only)

Area of Science		All	NAEP	non-NAEP
General Science	$\bar{\Delta}$	.29		.29
	s	.45		.45
	n	10		10
Biology	$\bar{\Delta}$	.02	.07	.01
	s	.15	.02	.17
	n	13	3	10
Earth Science	$\bar{\Delta}$	.07	.07	
	s	.17	.17	
	n	3	3	
"Physical Science"	$\bar{\Delta}$	.33	.18	.41
	s	.35	.05	.44
	n	8	3	5
Chemistry	$\bar{\Delta}$	.16		.16
	s	.29		.29
	n	8		8
Physics	$\bar{\Delta}$	.22		.22
	s	.12		.12
	n	3		3

of a standard deviation. Thus, with respect to important stepping stones to post-secondary programs and ultimately careers in science, women are at a disadvantage -- not because of statistically significant differences between the sexes, but rather because of the educationally significant magnitudes of those differences.

Lessons from Mathematics Education

Having accumulated a great deal of information on the nature of sex-related differences in science, researchers will want to shift their focus to explaining these differences. At least in terms of "time on task" mathematics educators are probably a few years ahead of science educators in this regard. During the latter part of the 1970s, many research studies, including some very large, federally funded projects, investigated the factors which seemed to be associated with the lower performance and participation of women in mathematics. While federal dollars for such research in science is not now so readily available (some such studies were funded), let us hope that the research efforts in this area do not "recreate the wheel". It is most likely that the causal factors perpetuating sex-related differences in the two fields are quite similar. In fact the two problems are even more closely tied because of the dependence of a great deal of scientific activity on mathematical competencies.

Of greater concern than covering the same ground is the possibility that researchers in science education will make the same mistakes the mathematics educators have made. Using research to explain group differences is far more difficult than using research to describe them. Kahl (1982) discussed the misleading and unjustified conclusions drawn by researchers hoping to explain sex-related differences in mathematics. Such conclusions have resulted from faulty assumptions, biased instruments and the misuse of various statistical techniques.

Earlier studies in mathematics placed undue emphasis on sex-related differences in suspected causal factors (Kahl, 1982). They overlooked the fact that differences between the sexes can exist in variables which are not at all related to mathematics performance and participation. They also failed to recognize that variables "operate differently" for different groups. One factor may be important for one group and not another. In such a case, does a group difference with respect to that variable matter? A sex-related difference in a variable is neither necessary nor sufficient

for that variable to contribute to an overall difference in some outcome.

While a great deal of attention has been given in recent years to test-item bias in cognitive tests, the same concern has not been directed toward researcher-made instruments, particularly in the affective domain (Kahl, 1982). For example, during test development, final test items have been selected which maximize the differences between the sexes. For some uses, the resulting test might be acceptable, but as a measure of the magnitude of a sex-related difference it would be biased. The use of the selection criterion would constitute "rigging the results". In such a case, the relationship between test usage and validity overrides the concern for construct validity in terms of consistency with previous theory or research. Would anyone advocate the same item-selection criterion for developing a cognitive test intended to be used to determine the magnitude of a sex-related difference?

Many statistical techniques have been applied which control for certain ariables in such a way that group differences in some outcome variable are reduced or even eliminated. Interestingly, when the same variables are controlled experimentally, the group differences often remain. The difficulty can often be traced to faulty assumptions about how variables operate for different groups (Kahl, 1982).

Multiple regression has been used frequently to determine the relative importance of predictor variables by attributing portions of variance in mathematics outcomes to various predictors and by attributing portions of a group difference in an outcome to corresponding group differences in a few predictors. Unfortunately, multiple regression doles out credit for variance in the dependent variable to predictors based purely on mathematical considerations, not some knowledge of reality. This makes absolutely no difference when multiple regression is used for purposes of prediction. However, the use of this analytic technique for explanation is highly overrated when comparisons of beta coefficients or related statistics are involved. It has led to too many misleading conclusions about the relative importance of possible causes of group differences (Kahl, 1982). In a landmark article, Darlington (1970) stated, "It would be better to simply concede that the notion of 'independent contribution to variance' has no meaning when predictor variables are intercorrelated."

These are some of the problems with research intended to explain sex-related differences in mathematics. There is no need to make the same mistakes in future research in other areas. It will continue to be very

important to monitor sex-related differences in cognitive and affective variables as well as in participation in various fields of study and work. Furthermore, few people would object to studies intended to provide explanations of sex-related differences as long as they were done well. Researchers in this area, however, cannot ignore the limitations of the methods available to them as is so often done.

The matter of research utility is also an important consideration. It can be argued that a better understanding of the causes of sex-related differences might influence the nature of interventions intended to alleviate the problem. However, in mathematics, the legitimate conclusions of research have only served to confirm what would logically occur to anyone with no knowledge of research results. Namely, if increased performance and participation of women in an area is desired, it would require direct encouragement by parents and teachers aimed at (1) improving the relevant attitudes of female students at all levels, (2) enhancing their self-confidence in the subject area and (3) increasing their awareness of the usefulness of the appropriate training in terms of keeping academic and career options open. Interestingly, success in these endeavors might in the long run influence those other "less intervenable" factors which receive so much attention in the literature.

REFERENCES

Darlington, R. B. Multiple regression in psychological research and practice. In E. Heermann & L. Braskamp (Eds.), Readings in statistics for the behavioral sciences. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1970.

Fleming, M. L. & Malone, M. A meta-analysis of the relationships between student characteristics and student outcomes in science. Draft of report on component of the NSF Science Meta-Analysis Project, University of Colorado, 1982.

Glass, G. V. Meta-analysis: An approach to the synthesis of research results. Journal of Research in Science Teaching, 1982, 19, 93-112.

Glass, G. V, McGaw, B. & Smith, M. L. Meta-analysis in social research. Beverly Hills, California: Sage Publications, 1981.

Kahl, S. R. Beyond "Women and mathematics: Does research matter?" Manuscript submitted for publication, March 1982.